****

**School of computer application**

CA Fourth

of

Big Data (CAP 457)

**Session 2022-2024**

**Submitted to: Submitted by:**

Name: Mr. Ravinder Singh Name: Satyam Mishra

(Dept of. Computer Science) Reg. no: 12212256

Course: MCA

**Department of computer Science**

**Lovely Professional University Jalandhar Punjab (144401)**

**India.**

1).Create database in mongo dB and create a collection in database.

Ans:-

Mongo DB is no-sql database which store data in a collection format which means in object format which have key and value pair

It is easy to use to use first we have to download the mongodb into system then configure it

To create collection first see how many databases you have using show dbs command it will list all database in mongodb

Now use which database you want to create collection using use and then database name.

Now inside that first create a collection name using db.createCollection(“collectionName”) it will create collection

Now inside that collection insert your data using db.collectionName.insert({“name”:”data”}) now it will create collection in your database

We can check it using command db.collectionName.find() it will show all collection of that database.

Now below are visual command to implement which I written upper.

Text

Description automatically generated

Here we first inside local database we create a collection name employee and inside that we inserted 2 collection and in last we able to see those 2 records.

2).demontarte the steps to set the environmental variable path for mongoDb.

Ans:-

To use mongoDB into your system first we have to download exe file of mongo DB into your system then just install that

After installation the file of mongo DB is stored into program file

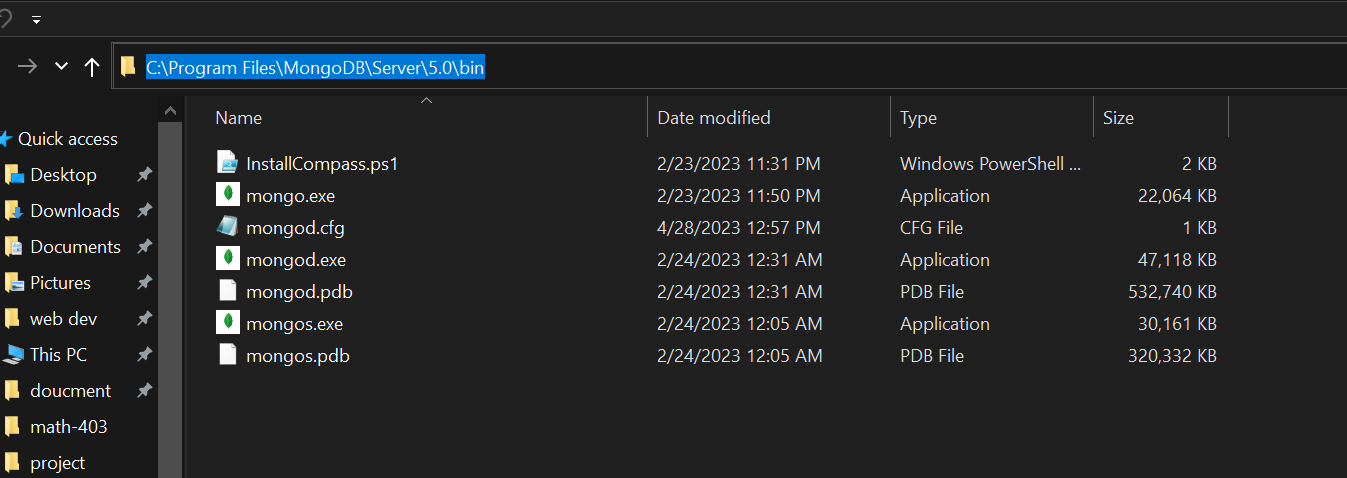
Now go to c drive and then program file then inside mongodb/server/5.0/bin and copy this path

Now go to search bar and search environment variable and open that now click on environment variable and click on path and add a new path with which you copy from program file mongo db bin folder and click ok then again ok

Now to check mongo dB perfectly configure or not go to cmd type mongo if mongo shell opening then mongo dB perfectly working on your system

Now do whatever you want .

Example : -



A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer error

Description automatically generated with medium confidence

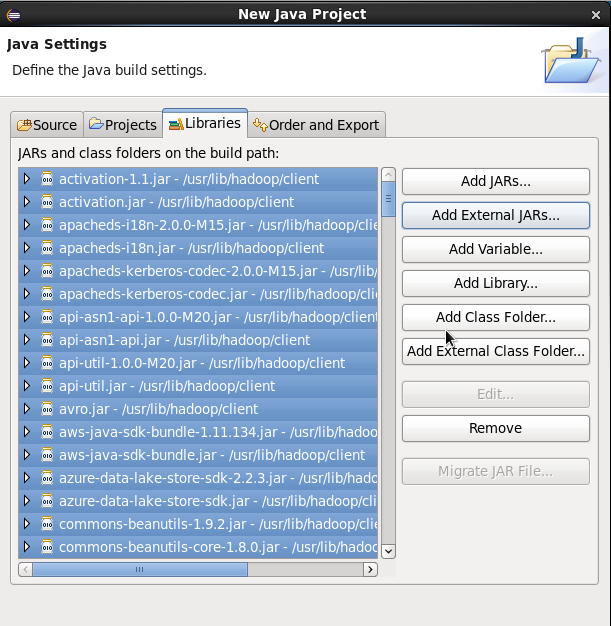
Q3. Discuss the concept of the map Reduce with the help of suitable examples.

Ans:-

**MapReduce is a software framework and programming model used for processing huge amounts of data. MapReduce program work in two phases, namely, Map and Reduce. Map tasks deal with splitting and mapping of data while Reduce tasks shuffle and reduce the data.**

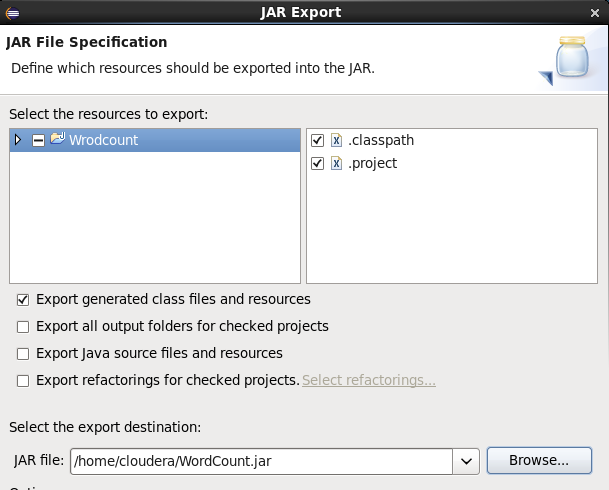
* **Example of Wordcount program using MapReduce.**
* **For Creating word count program in Hadoop Follow the steps:**
* Graphical user interface

  Description automatically generated**Open Eclipse**
* **Create a new java project name it wordcount.**

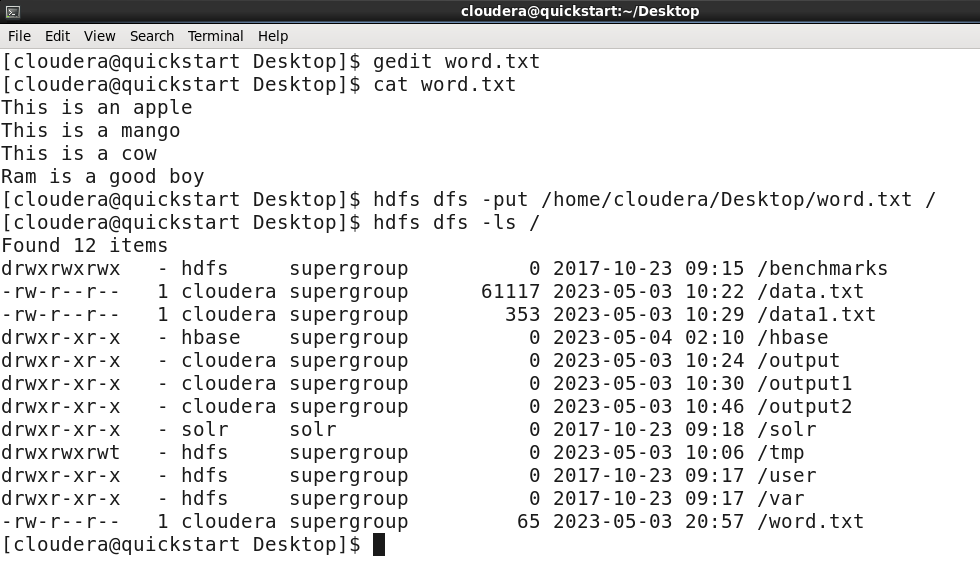


* **Then click on the “src” and create a new class name it “WordCount”**
* **Then write the java code for the wrodcount**
* **Save it**
* A screenshot of a computer

  Description automatically generated with medium confidence**Then right click on the wordcount project file and export it as JAR and save this jar file at any location you want I am saving it in “/home/cloudera/Wordcount.jar”**



* **Now create a txt file in local (Linux) environment and fill in some text/sentences in it. In this txt file the wordcount operation is performed.**
* **Put this file into Hadoop directory.**



* **Now type the command:**

**Hadoop jar /home/cloudera/WordCount.jar /word.txt /process**

* + **/home/cloudera/WordCount.jar: - JAR file location**
  + **/word.txt – txt file on which wordcount operation is to be perform**
  + **/process – The directory where the output result will be store**
* A screenshot of a computer

  Description automatically generated**After executing this command the output will store in /process directory in Hadoop with the file name part-r-00000 You can see the content using cat command.**

